



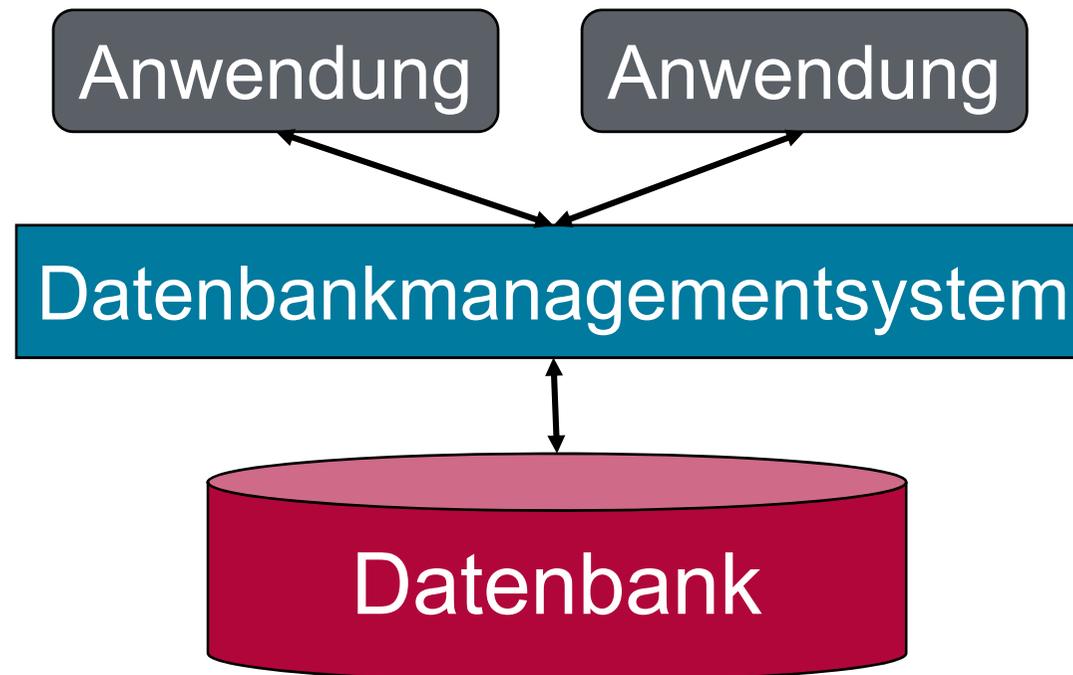
**Hasso
Plattner
Institut**

IT Systems Engineering | Universität Potsdam

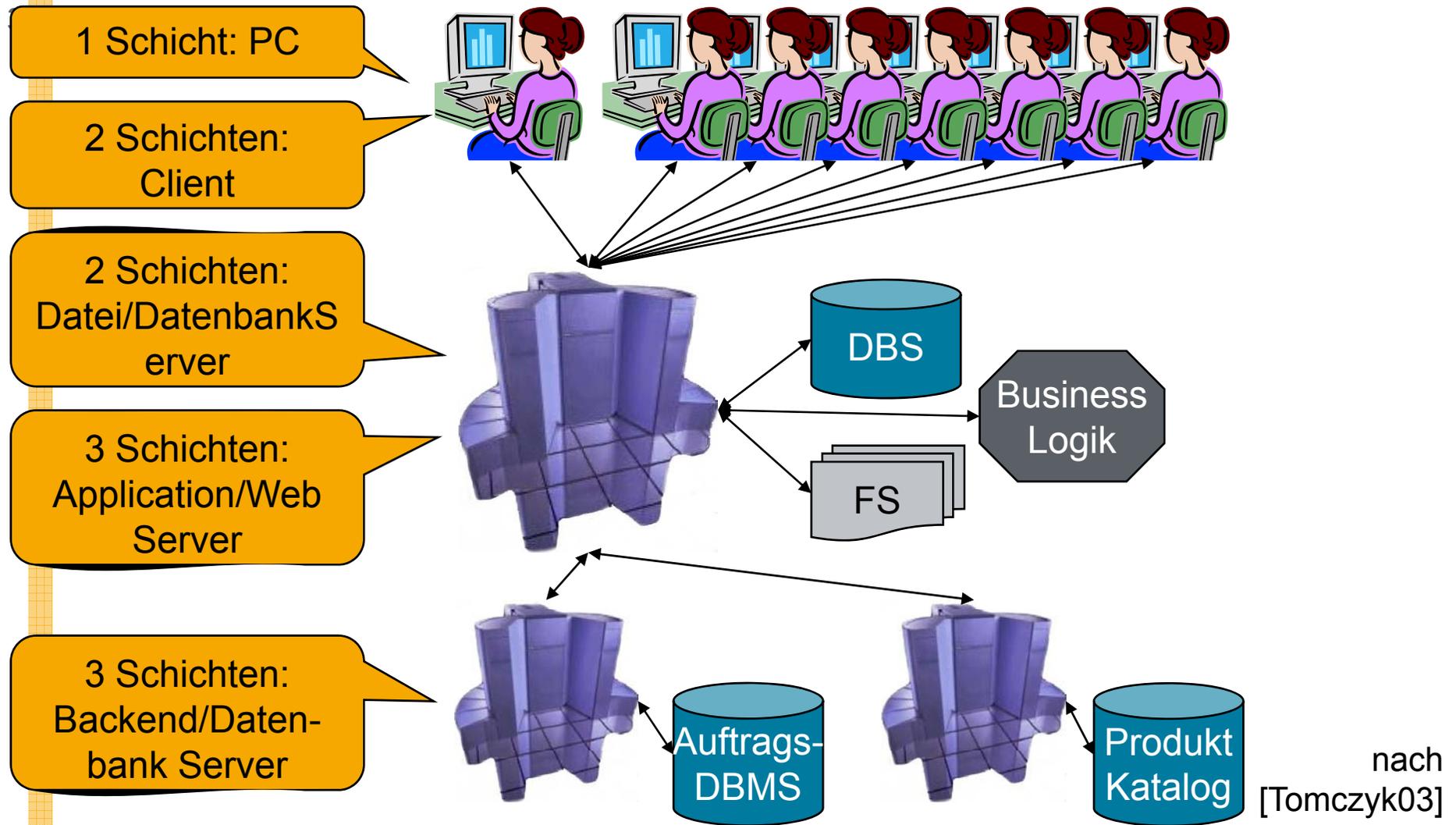
Datenbanksysteme 1
Organisatorisches und Einführung

15.10.2007

Felix Naumann

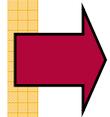


Mehr-Schichtenarchitekturen



nach [Tomczyk03]

4



- Vorstellung der Arbeitsgruppe
- Organisatorisches
- Datenbanken und Informationssysteme
- Das Semester an einem Beispiel
- Ausblick auf das Semester



- Felix Naumann
 - naumann@hpi.uni-potsdam.de
 - Raum A-1.13 (über Frau Pamperin)
 - HU, IBM Almaden, DFG/JP/HU, HPI
- Universität Potsdam, Hasso-Plattner-Institut
 - Fachgebiet „Informationssysteme“
- Forschungsthemen (gleich mehr):
 - Datenqualität
 - Informationsintegration
 - Peer Data Management
 - Search
 - <http://www.hpi.uni-potsdam.de/naumann/>

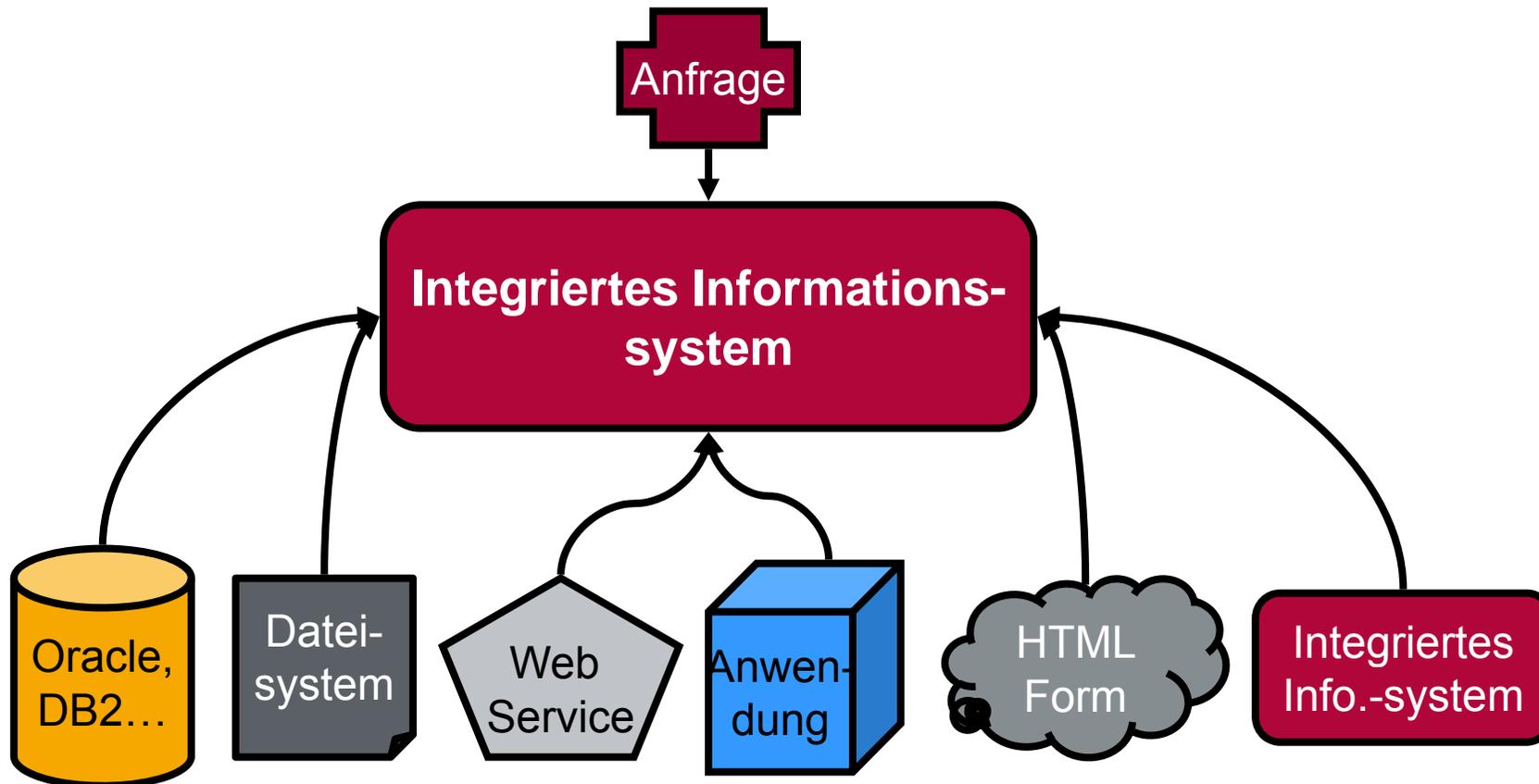
Die Arbeitsgruppe Informationssysteme



6

- Wissenschaftliche Mitarbeiter
 - Alexander Albrecht: ETL, Schema Mapping, PIM
 - Jana Bauckmann: Data Profiling, Aladin
 - Jens Bleiholder: Data Fusion, HumMer & FuSem
 - Paul Führung: DQ Assessment, Viqtor
 - Frank Kaufer: Schema and Ontology Matching
 - Armin Roth: Peer-Data-Management, System P
 - Melanie Weis: Duplicate Detection, DogmatiX
- Studentische Hilfskräfte
 - Karsten Draba: HumMer & FuSem
 - N.N: Ranking, SPRINT
 - Tobias Flach: Aladin Project
 - N.N.: System P
- <http://www.hpi.uni-potsdam.de/naumann/>





Schematische und Daten-Heterogenität

9

Variante 1

Männer	
Vorname	Nachname
Felix	Naumann
Jens	Bleiholder

Frauen

Vorname	Nachname
Melanie	Weis
Jana	Bauckmann

Variante 2

Personen			
Vorname	Nachname	Männl.	Weibl.
Felix	Naumann	Ja	Nein
Jens	Bleiholder	Ja	Nein
Melanie	Weis	Nein	Ja
Jana	Bauckmann	Nein	Ja

Variante 3

Personen		
Vorname	Nachname	Geschlecht
Felix	Naumann	Männlich
Jens	Bleiholder	Männlich
Melanie	Weis	Weiblich
Jana	Bauckmann	Weiblich

Schematische und Daten-Heterogenität

10

Variante 1

Männer	
Vorname	Nachname
Felix	Naumann
Jens	Bleiholder

Frauen

Vorname	Nachname
Melanie	Weis
Jana	Bauckmann

Variante 2

Personen			
FirstNa	Name	male	femal
Felix	Naumann	Ja	Nein
Jnes	Bleiho.	Ja	Nein
Melanie	Weiß	Nein	Ja
Jana	baukman	Nein	Ja

Variante 3

Personen		
VN	NN	SEX
F.	Naumann	Männlich
J.	Bleiholder	Männlich
M.	Weis	Weiblich
J.	Bauckmann	Weiblich

Schematische und Daten-Heterogenität

11

Variante 1

*■□■□■		□□□■□■	
†□□□■	♣■□□□■	†□□□■	♣■□□□■
♣■●×	♣■◆○□■	♣■●●	♣■×◆
☺■□◆	×■□■●□■□	☺■□	♣■&○□■

Variante 2

♣■□◆□■□■			
†□□■	♣■□□■	■●□	●□
♣■●×	♣■◆○□■	□	□
☺■□◆	♣■●□■	□	□
♣■●●	♣■×◆	□	□
☺■□	&○□■	□	□

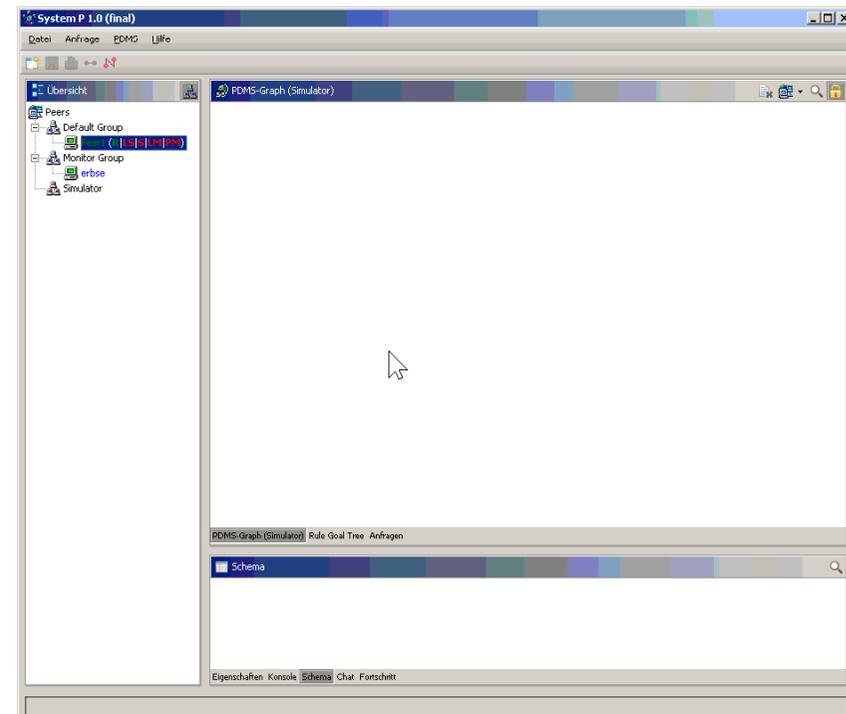
Variante 3

♣■□◆□■□■		
■□□■	♣■□□■	♣■□■◆
♣■●×	♣■◆○□■	♣■●×■
☺■□◆	♣■●□■	♣■●×■
♣■●●	♣■×◆	♣■●×■
☺■□	♣■◆■&	♣■●×■

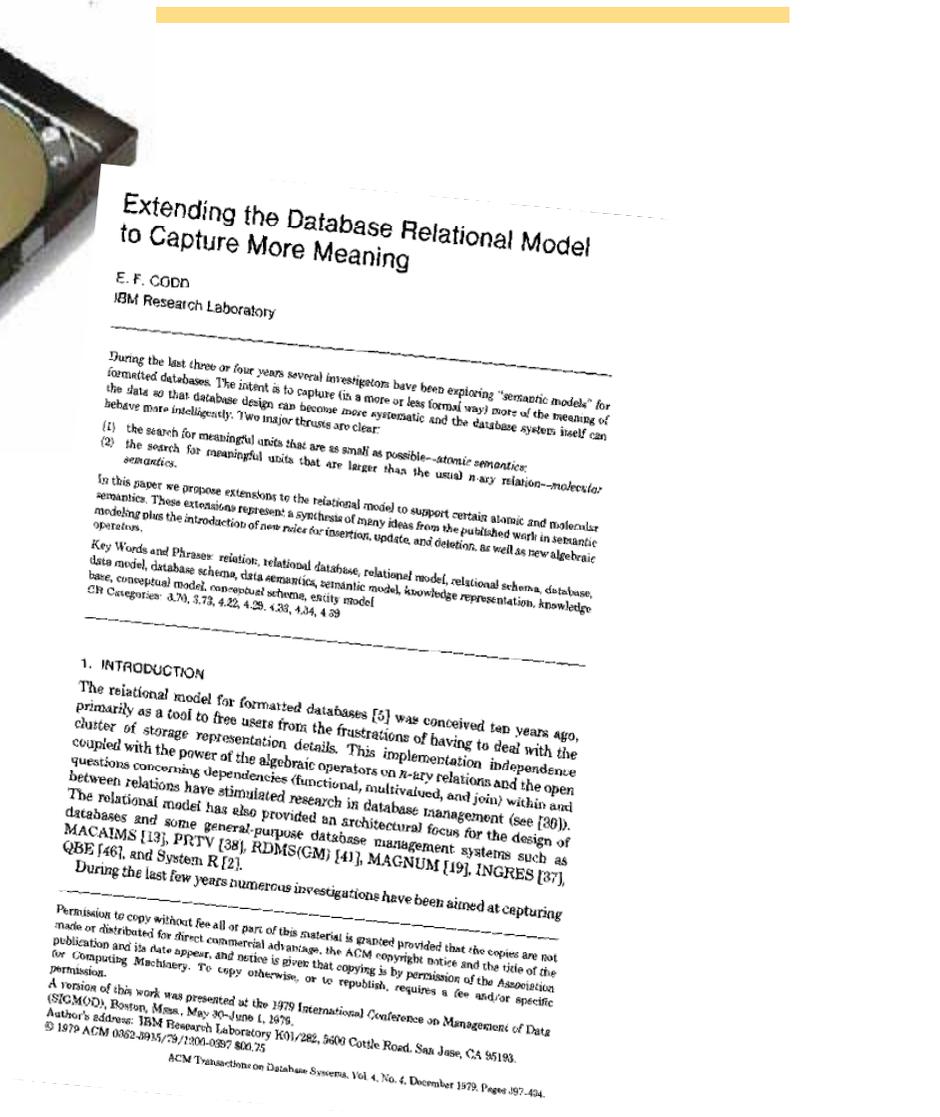
Weitere Forschungsthemen

12

- Informationsintegration
 - Schema Matching
 - Duplikaterkennung
 - Datenfusion
- Datenqualität
- Peer Data Management
- Life Sciences: Aladin
- Search



- Vorlesungen
 - DBS I (2. Semester)
 - Ab SS 2008
 - DBS II (3. Semester)
 - Ab SS 2008
 - Informationsintegration (Master)
 - ...
- Seminare
 - Beauty is our business
 - www.ProminentPeople.info
 - Datenfusion in drei Schritten
 - Schema Matching

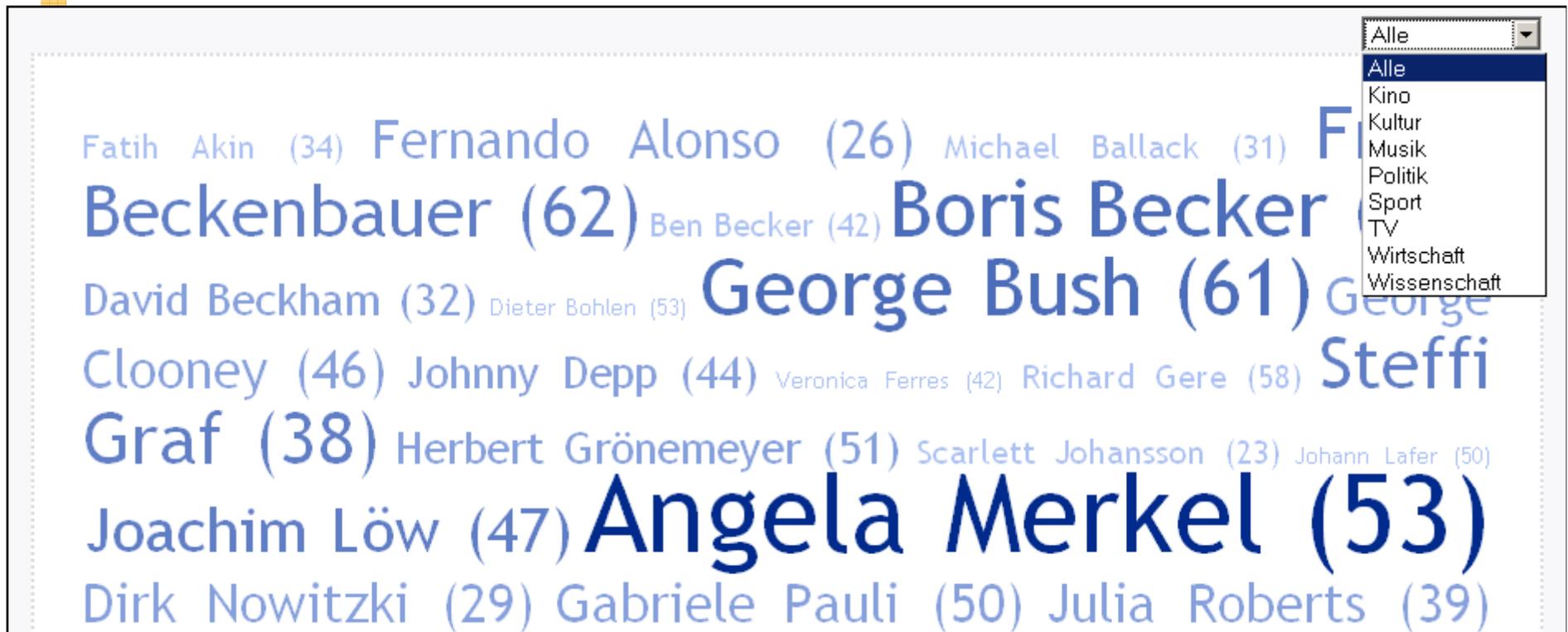


Beauty is our Business

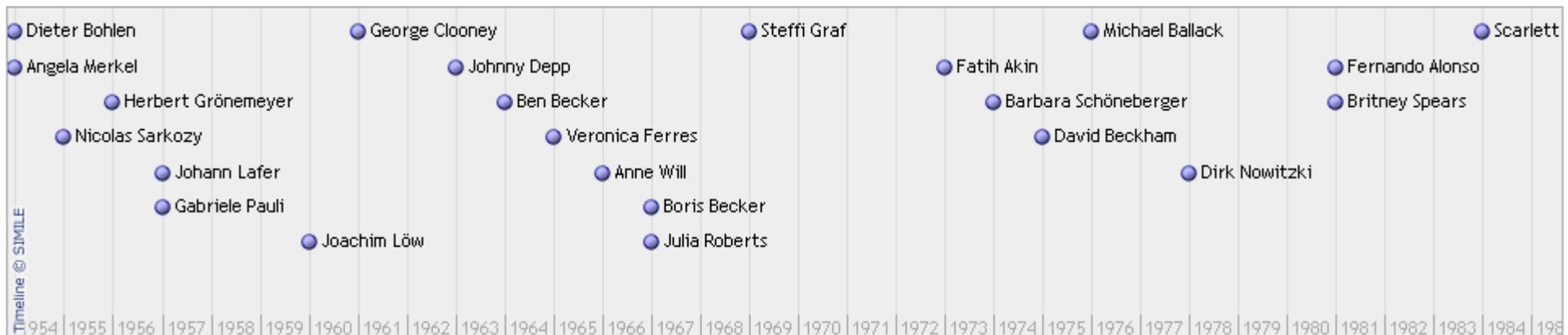
„Wenn wir uns klarmachen, daß der Kampf gegen Chaos, Durcheinander und unbeherrschte Kompliziertheit eine der größten Herausforderungen der Informatik ist, müssen wir zugestehen: Beauty is our Business.“ Edsger W. Dijkstra, 1978

Termin	Thema	Vortrag
22.10.2007	Einführung	Felix Naumann
t.b.d.	Wissenschaftliche Texte Lesen	Felix Naumann
12.11.2007	Literaturkritik / Diskussion	Alle
19.11.2007	Vortragstechniken	Felix Naumann
3.12.2007	Vortrag 1: Mariposa (pdf) Vortrag 2: Trio (pdf)	
10.12.2007	Einführung in LaTeX	Felix Naumann
17.12.2007	Vortrag 3: Fagins Algorithmus (pdf) Vortrag 4: Enough Already in SQL (pdf)	
7.1.2008	Vortrag 5: Sorted Neighborhood (pdf) Vortrag 6: Data Mining (pdf)	
14.1.2008	Vorstellung der Gliederungen & Tipps zur Ausarbeitung	Alle
29.2.2008	Abgabe der Ausarbeitungen	Alle

- Extraktion von Prominenten (Alter, Titel) aus News Texten – **N**amed **E**ntity **R**ecognition (NER)
- Visualisierung der Ergebnisse auf www.ProminentPeople.info als Name Cloud



- Datenanalyse, insbesondere für Altersangaben
 - Anfrage nach Alter (mit Widersprüchen)
 - Anfrage nach Geburtstag (auch aus statistischer Analyse)
 - Visualisierung über Zeitleiste, <http://simile.mit.edu/timeline/>



Vorstellung – Hörer

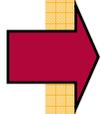
17

- Welches Semester?
- HPI oder IfI?
- Erasmus o.ä.?
- Berlin/Brandenburg oder anderswo?

- Berührung mit Datenbanken?
 - Wie?

Überblick

18



- Vorstellung der Arbeitsgruppe
- Organisatorisches
- Datenbanken und Informationssysteme
- Das Semester an einem Beispiel
- Ausblick auf das Semester



Termine

19

- Vorlesung
 - Montags 13:30 – 15:00
– Hörsaal 2
 - Mittwochs 13:30 – 15:00
– Hörsaal 3
- Erste Vorlesung
 - 15.10.2007
- Letzte Vorlesung
 - 6.2.2008
- Weihnachten
 - 24.12.2006 – 6.1.2007
- Feiertage
 - Reformationstag am 31.10.
- Übungen
 - Ungefähr 1 SWS
 - Verteilt auf Doppelstunden im Semester
 - Zwei parallele Übungsgruppen
 - Erste Übung
– 29.10.2007
 - Übungsleitung: Jana Bauckmann
- Klausur
 - Erste Woche nach Semester

- Folien
 - Vor der VL im WWW
 - ASAP
 - Datiert wegen updates
- Anregungen zur Verbesserung:
 - Gebrauch der Folien
 - Infos im WWW
 - Übungsbetrieb
 - ...
- Fragen bitte jederzeit!
 - In der Vorlesung
 - Sprechstunde
 - Dienstags 15:00 – 16:00
 - Raum A-1.13
 - Am liebsten mit Anmeldung
 - Email: naumann@hpi.uni-potsdam.de

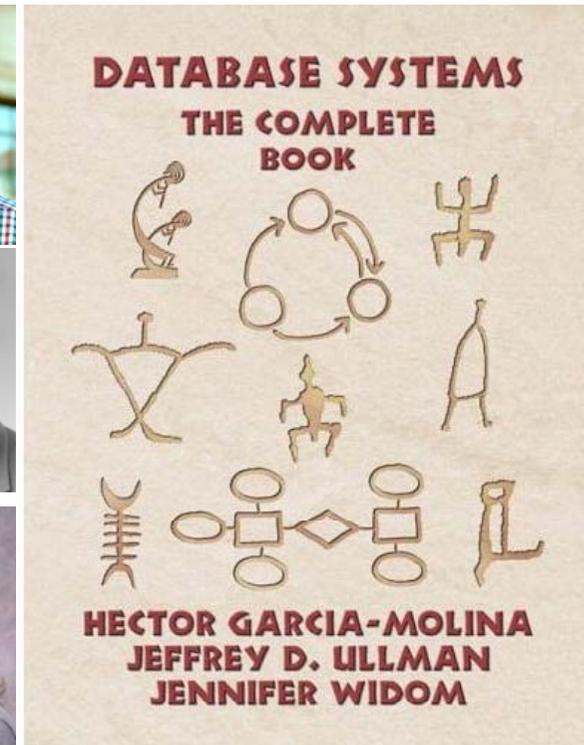
Database Systems - The Complete Book

- Hector Garcia-Molina, Jeffrey D. Ullman, Jennifer Widom: , Pearson Education International, 2002.
- Viele Exemplare in Bibliothek
- Auch für DBS II !

Ebenfalls empfehlenswert

- Datenbanken Konzepte und Sprachen
Andreas Heuer, Gunter Saake
ISBN: 3826606191
- Grundlagen von Datenbanksystemen.
Ausgabe Grundstudium.
Ramiz Elmasri, Shamkant B. Navathe

Und viele andere mehr...



Übungen und Leistungserfassung

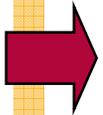
22

- Übungen – Organisatorisches
 - Zweiergruppen
 - Mind. 50% der Punkte auf jedem Blatt
 - Ein Freischuss
 - Tutoren
 - Lutz Gericke
 - Mehr von Jana Bauckmann bei erster Übung
- Übungen – Inhaltliches
 - „Theoretische“ Übungen
 - Auf Papier
 - Praktische Übungen
 - DB2 im Studentenpool
 - Ergebnisse auf Papier
- Klausur
 - 90 Minuten

Überblick

23

- Vorstellung der Arbeitsgruppe
- Organisatorisches
- Datenbanken und Informationssysteme
- Das Semester an einem Beispiel
- Ausblick auf das Semester



Was sind Daten?

24

Digitale Repräsentation von

- Dingen
- Entitäten
- Wissen
- Information

in der wirklichen Welt.

Kernfragen:

- Welche Daten speichere ich?
- Wie speichere ich die Daten?
- Wie frage ich Daten ab?
- Wie erledige ich all dies effizient und sicher?

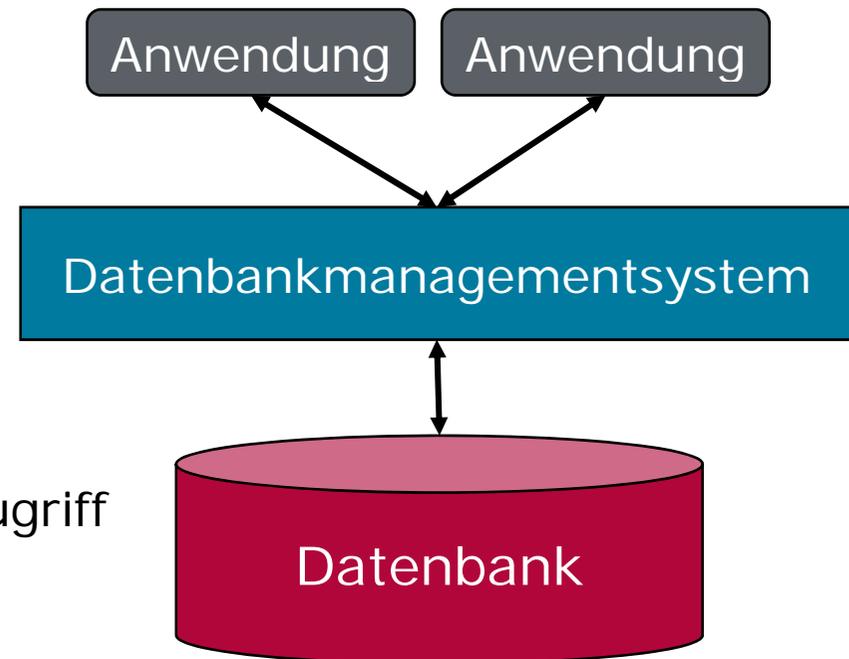
} Unterstützung durch
Datenbanksystem

Datenbanksysteme

25

Bestandteile

- Datenbank
 - Die Daten selbst
 - + Metadaten (Beschreibung der Daten)
- Datenbankmanagementsystem (DBMS)
 - Softwarekomponente zum Zugriff auf eine oder mehrere Datenbanken.
 - Server-basiert
- (Anwendungen)



DBMS Beispiele

26

DBMS

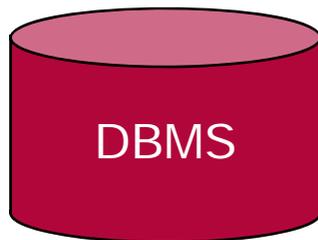
- OLTP (Online Transaction Processing)
- Oracle, DB2, SQL-Server
- Informix, Sybase
- Teradata
- PostGreSQL, Interbase
- mySQL, ...

Data Warehouses

- OLAP (Online Analytical Processing)

Alle großen Softwaresysteme nutzen ein, mehrere oder sehr viele DBMS.

- SAP
- Siebel
- SABRE
- Sogar (zwischenzeitlich) das geplante Microsoft Windows File-System (Longhorn)



ORACLE

tamino

SYBASE

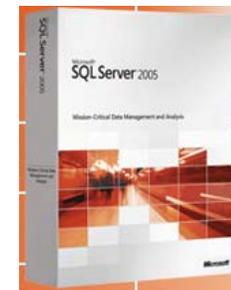


DB2 9 for Linux UNIX and Windows

pureXML™ and storage compression

Felix Naumann | VL Datenbanksysteme I | WS 07/08

Information Management
software



DBMS Aufgaben

27

- Unterstützung des Datenmodells
- Bereitstellung einer Anfragesprache (DDL & DML)
- Effiziente Anfragebearbeitung
- Robustheit
 - Wahrung der Datenintegrität (Konsistenz etc.)
 - Abfangen von Systemfehlern
- Speicherverwaltung (RAM & Disk)
- Transaktionsmanagement
 - Auch im Mehr-Benutzer-Betrieb
- Nutzerverwaltung & Zugangskontrolle

Beispiele für Informationssysteme

28

Dateisystem

- Informationseinheit:
Flat file
- Anfrage:
File search, RegEx
- Struktur:
Flach oder hierarchisch

Beispiele

- NTFS
- NIS
- FTP Zugriff

Einsatzgebiete

- WWW (HTML Dateien)
- Desktop-Anwendungen
(Textverarbeitung, etc.)

Beispiele für Informationssysteme

29

Datei

- Informationseinheit:
Zeile /Token

- Anfrage:
Parser

- Struktur:
Flach

Beispiele

- Komma-delimited files
- Annotated files

Einsatzgebiete

- SwissProt

```

ID  RNGTPCHI  standard; RNA; ROD; 1016 BP.
XX
DT  01-AUG-1991 (Rel. 28, Created)
DT  04-MAR-2000 (Rel. 63, Last updated, Version 2)
XX
DE  Rat GTP cyclohydrolase I mRNA, complete cds.
XX
KW  GTP cyclohydrolase I.
XX
OS  Rattus norvegicus (Norway rat)
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC  Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
XX
RN  [1]
RP  1-1016
RX  MEDLINE; 91093270.
RX  PUBMED; 1985963.
RA  Hatakeyama K., Inoue Y., Harada T., Kagamiyama H.;
RT  "Cloning and sequencing of cDNA encoding rat GTP cyclohydrolase I: The
RT  first enzyme of the tetrahydrobiopterin biosynthetic pathway";
RL  J. Biol. Chem. 266(2):765-769(1991).
XX
FT  CDS          128..853
FT              /codon_start=1
FT              /db_xref="GOA:P22288"
FT              /db_xref="SWISS-PROT:P22288"
FT              /EC_number="3.5.4.16"
FT              /gene="GTP cyclohydrolase I"
FT              /product="GTP cyclohydrolase I"
FT              /protein_id="AAA41299.1"
FT              /translation="MEKPRGVRCNTNGFPERELPRPGASRPAEKSRRPEAKGAQPADAWK
FT              AGRPRSEEDMELNLPNLAAYSSILRSLGEDPQRQGLLKTFWRAATAMQFFTRGYQETI
FT              SDVLNDAIFDEHDDEMIVKIDIMFMSCEHHLVFPVGRVHIGYLPNKQVLGSLKLARIV
FT              EIYSRRLQVQERLTKQIAVAITEALQPAAGVGVVIEATHMCMVMRQKMNKSTVTSML
FT              GVFREDPKTREEEFLTLIRS"
SQ  Sequence 1016 BP; 236 A; 279 C; 291 G; 210 T; 0 other;
gacttcgaac ctcattcggg gcagaactcc tgtcccgggtg acagccacag gtcacggcgc      60
cgggctaagc cgagccgcag cgcttggttag caccttaggg tgtctegga gcaatcgcg      120
cgggtccatg gagaagccgc ggggtgtaag gtcaccaat gggttcccc agcgggagct      180
...
catcaggagc tgaacttccg tgtgcgagcc cgggtttgca gacccccgct gaggccagcg      900
ttatctgtct cgattgtaca ttccagttcc agttggtata cttgtcaact ttatttctca      960
ccatgaattg tattaataa ttatttatag agatgtcaaa taaaggtgat caactt      1016
//

```

Molecule type Name

Date of creation and last update

Free text description

Keywords describing the molecule

Organism

Article the sequence was published in

Structural annotation (coding sequence)

Link to functional annotation of resulting protein

Translated protein sequence

Sequence of bases

Beispiele für Informationssysteme

30

Markup Datei

- Informationseinheit:
Tagged text
- Anfrage:
Parser, Anfragesprache
- Struktur:
Flach, hierarchisch oder graph-
basiert

Beispiele

- XML
- HTML

Einsatzgebiete

- Web Services
- Messages
- Interoperationale Anwendungen

Beispiel: XML-Dokument

31

```
<?xml version="1.0" encoding="UTF-8"?>
<rechnung kundennummer="k333063143">
  <monatspreis>0,00</monatspreis>
  <einzelverbindungsnaechweis>
    <verbindung>
      <datum>26.2.</datum>
      <zeit>19:47</zeit>
      <nummer>200xxxx</nummer>
      <einzelpreis waehrung="Euro">0,66</einzelpreis>
    </verbindung>
    <verbindung>
      <datum>27.2.</datum>
      <zeit>19:06</zeit>
      <nummer>200xxxx</nummer>
      <einzelpreis waehrung="Euro">0.46</einzelpreis>
    </verbindung>
    <verbindungskosten_gesamt waehrung="Euro">2.19</verbindungskosten_gesamt>
  </einzelverbindungsnaechweis>
</rechnung>
```

Beispiele für Informationssysteme

32

Datenbank

- Anfrage:
Komplexe Sprache, z.B. SQL
- Informationseinheit:
Tupel / Attribut, Objekt
- Struktur:
Relational, OO, Hierarchisch

Beispiele

- Relationale DBMS
- OO DBMS
- Auch: XML DBMS

Einsatzgebiete

- Data Warehouses
- OLTP
- Banken/Versicherungen

Beispiele für Informationssysteme

33

HTML Formular

- Informationseinheit:
HTML Seite, Text
- Anfrage:
Suchworte, Formular (inkl.
Radiobutton, dropdown-list, etc.)
- Struktur:
wie Markup Datei: Flach,
hierarchisch oder graph-basiert
I.d.R. flach

Beispiele

- Einfache Suchformulare
- Komplexe Anfrageformulare

Einsatzgebiete

- Suchmaschinen
- Reisedienste
- Kataloge

Dahinter: Fast immer ein DBMS!

Amazon.de Bücher: Erweiterte Suche - Mozilla Firefox

http://www.amazon.de/Erweiterte-Suche-B%C3%BCcher/b/ref=sv_b_0/302-0722786-8511230?ie=

LEO de<->en

Groups Google DBLP eBay a Dilbert MAC RPI FGIS Wiki QDB 2007 Lehre G Scholar Maps RPI TYPO3

Amazon.de Bücher: Erweiterte Su...

amazon.de

WUNSCHZETTEL | MEIN KONTO | HILFE | IMPRESSUM

HOME FELIX'S SHOP **BÜCHER** ENGLISH BOOKS ELEKTRONIK & FOTO MUSIK DVD KAUFEN & LEIHEN SOFTWARE GAMES KÜCHE, HAUS & GARTEN SPIELWAREN & KINDERWELT SPORT & FREIZEIT UHREN BABY **NEU** SCHUHE & HANDTASCHEN

ERWEITERTE SUCHE | STÖBERN | BESTSELLER | NEUHEITEN | HÖRBÜCHER | TASCHENBÜCHER | FACHBÜCHER | PREIS-HITS | BÜCHER VERKAUFEN

Suche Bücher LOS

HARRY POTTER und die Heiligtümer des Todes  Frühstück mit Harry [Jetzt portofrei vorbestellen](#) am 27.10.2007 bis 10:30 Uhr geliefert

Erweiterte Suche Bücher

Je mehr Felder Sie ausfüllen, desto zielgerichteter können wir suchen. Es reicht jedoch aus, nur eines der Felder auszufüllen.

Autor/in:

Titel:

Schlagwörter:

ISBN: (10- oder 13-stellig, ohne Bindestriche)

Verlag:

Verfeinern Sie Ihre Suche, indem Sie nur nach bestimmten Buchformaten suchen lassen.

Nur gebraucht:

Format: Alle Formate

Ordnen nach: Topseller

Erscheinungsdatum: vor Jeder Monat 2009

Suche in: deutsche Bücher englische Bücher

Jetzt suchen

Beispiele

Search Books at Amazon.com - Mozilla

File Edit View Go Bookmarks Tools Window Help

Back Forward Stop Reload

http://www.amazon.com/exec/obidos/ats-query-page/ref=b_bh_lr_a_2/102-19 Search Print

Home Yahoo Groups Google DBLP Dilbert HiIQ Forms Berlin BVG MAC LaTeX CiteSeer

Shop in **Sporting Goods** (Beta-What is this?) **SPORTING GOODS**

amazon.com   [VIEW CART](#) | [WISH LIST](#) | [YOUR ACCOUNT](#) | [HELP](#)

[WELCOME](#) [FELIX'S STORE](#) **BOOKS** [APPAREL & ACCESSORIES](#) [ELECTRONICS](#) [TOYS & GAMES](#) [KITCHEN & HOUSEWARES](#) [SOFTWARE](#) [SEE MORE STORES](#)  **Felix's Gold Box**

[SEARCH](#) [BROWSE SUBJECTS](#) [BESTSELLERS](#) [MAGAZINES](#) [CORPORATE ACCOUNTS](#) [E-BOOKS & DOCS](#) [BARGAIN BOOKS](#) [USED BOOKS](#)

Search Books

Fill in **at least** one field. Fill in more to narrow your search. Need more flexibility? Try [Power Search](#). Need help? Go to [search tips](#).

Author: [Search Now](#)

Title:

Subject:

ISBN:

Publisher:

Refine your search (optional):

Used Only:

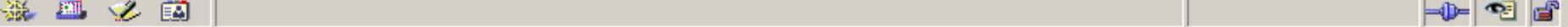
Format:

Reader age:

Language:

Publication date: (e.g. 1999)

Sort results by:

Taskbar: 

Beispiele für Informationssysteme

36

Web Service

- Informationseinheit:
XML Dokument
- Anfrage
XML Dokument
- Struktur:
Wie XML: Flach,
hierarchisch, graph-
basiert

Beispiele

- Einfach: Temperaturdienst,
etc
- Komplex: Reservierungen
(Schachtelung und
Verknüpfung von Web
Services)

Einsatzgebiete

- Intra-organisatorische
Workflows
- E-Marketplaces
- Datenaustausch

Dahinter: Fast immer ein DBMS!

www.xmethods.net - Mozilla Firefox

Datei Bearbeiten Ansicht Gehe Lesezeichen Extras Hilfe

http://www.xmethods.com/ve2/Directory.poj;sessionid=88iPof4Dr1Au5t5hV5CZrRZ5(QHyMHIRM)

Yahoo Groups Google DBLP eBay Amazon Dilbert Yahoo! MAC MACWiki Lehre Scholar Maps ManDup TYPO3 HPI

HPI: Datenbanksysteme I Computer reservations system - ... Article not found - Wikipedia, the f... www.xmethods.net



[Home](#) · [Interfaces](#) · [Tools](#) · [Implementations](#) · [Manage](#) · [Register](#) · [Tutorials](#) · [About](#)

Full Service List

Also accessible via XML Interfaces: [DISCO](#) [WS-Inspection](#) [RSS](#) See the [interfaces section](#) for more information.

List is ordered by submission time, with most recent services listed first.

Publisher	Style	Service Name	Description	Implementation
aandreu	RPC	Try It neuroFuzz Crypto Service	A simple demonstration of AES encryption and decryption via SOAP.	NuSOAP
XWebServices	DOC	Try It XWebBlog	XML/SOAP based Web Service which provides Content Management integration of a Weblog (Blog) to client applications.	MS .NET
StrikeIron	RPC	Try It LMSecurities	Latest stock quotes from Warsaw Stock Exchange	
kylehayes	RPC	Try It Amortization Calculator	Calculates amortization given principal, periodic interest rate, total number of payments	
jbardin	RPC	Try It Interconnect	Initiates calls between one or more phones	SOAP4R
StrikeIron	DOC	Try It U.S. Yellow Pages	Access to yellow pages listings for 17 million U.S. businesses	
7leapcom	RPC	Try It Conversão de IP para País	Converts an IP do it's country origin, retrieves not only country name, as well iso code, region, capital, and coin	NuSOAP
7leapcom	RPC	Try It Custo de Envio de Encomendas pelos CTT	Calculates shipping costs from Portugal via nacional CTT post offices	NuSOAP
agenteel	RPC	Try It Codigos Postales de Mexico	Codigos Postales de Mexico is a client/server implementation to lookup Postal Codes in Mexico.	gSOAP
Xignite	DOC	Try It XigniteCompensation	Biographical and Compensation information for more than 100,000 US executives.	MS .NET
Xignite	DOC	Try It XigniteWatchLists	Provide for the management of financial watchlists (stocks, currencies, rates...)	MS .NET
Xignite	DOC	Try It XigniteOFAC	Enables OFAC Compliance against the Specially Designed Nationals database	MS .NET

Beispiele für Informationssysteme

38

Anwendung

- Informationseinheit:
Java Objekt, Text
- Anfrage:
via Anwendungsschnittstelle oder GUI
- Struktur:
Objekt (Interface)
Display (GUI)

Beispiele

- Java, C++, etc
- Legacy Informationssysteme

Einsatzgebiete

- Komplexe Analysen (Data Mining, Statistik)

Dahinter: Fast immer ein DBMS!

Beispiele für Informationssysteme

39

Integriertes Informationssystem

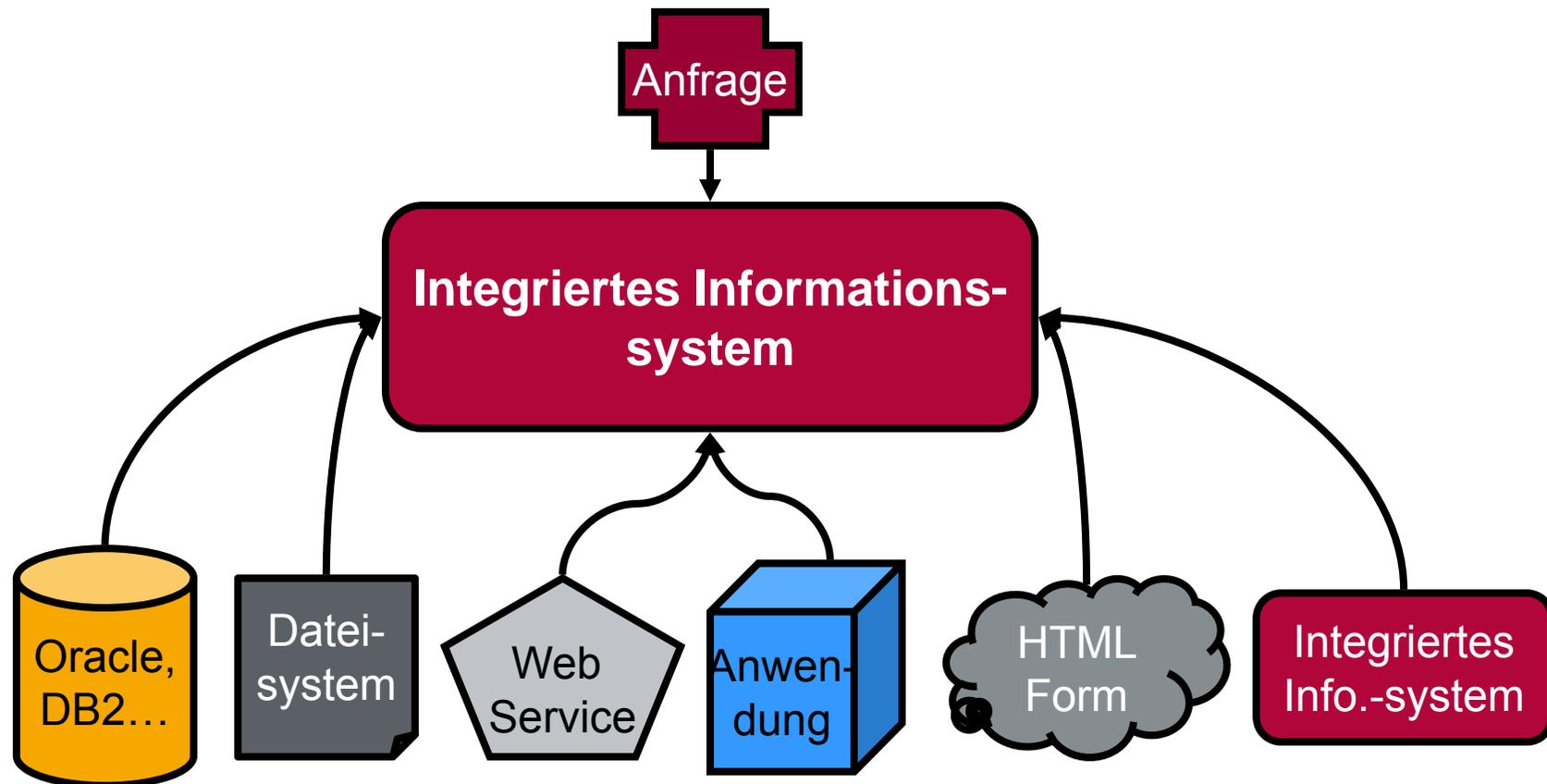
- Verhält sich in Anfrage, Struktur und Informationseinheit je nach Design:
 - DBMS
 - HTML Formular
 - Web Service
 - ...

Beispiele

- Data Warehouses
- Föderierte Datenbanken
- Portale

Einsatzgebiete:

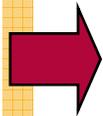
- Meta Search
- Life Sciences
- Int. Unternehmen
- Intranets



Überblick

41

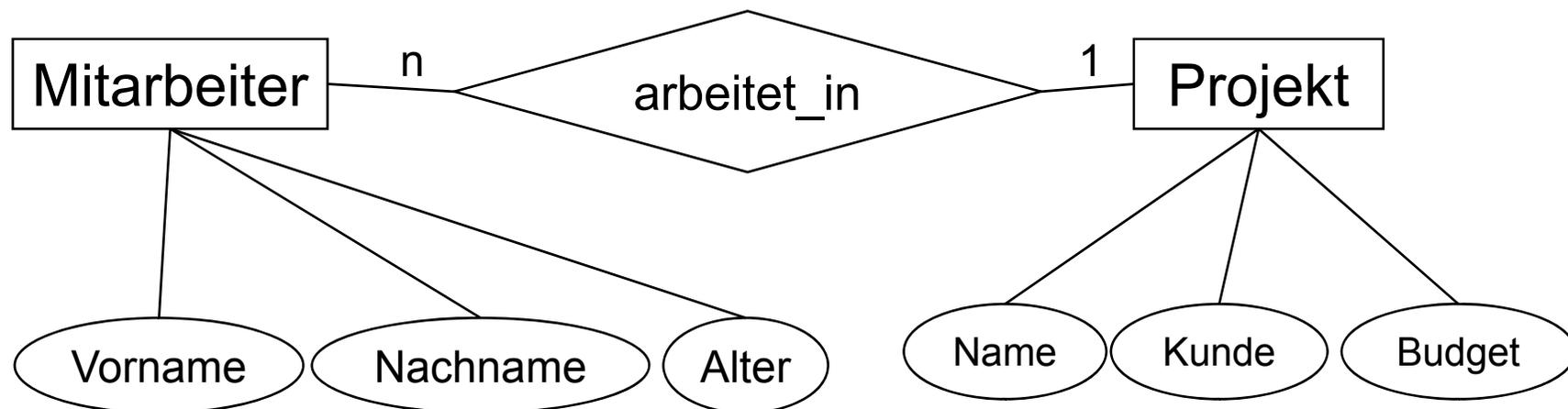
- Vorstellung der Arbeitsgruppe
- Organisatorisches
- Datenbanken und Informationssysteme
- Das Semester an einem Beispiel
- Ausblick auf das Semester



Datenbankentwurf

42

Mitarbeiter und Projekte



Relationales Modell – Relationen

43

mitarbeiter				
p_id	vorname	nachname	alter	proj_id

```
CREATE TABLE mitarbeiter(
    p_id          INTEGER,
    vorname       CHAR(25),
    nachname      CHAR(50),
    alter         INTEGER,
    proj_id       INTEGER,
    PRIMARY KEY(p_id),
    FOREIGN KEY proj_id
    REFERENCES projekte )
```

projekte			
proj_id	name	kunde	budget

Relationales Modell – Extension

44

mitarbeiter

p_id	vorname	nachname	alter	proj_id
1	Peter	Müller	32	1
2	Stefanie	Meier	34	1
5	Petra	Weger	28	2
7	Andreas	Zwickel	44	5
...	

projekte

proj_id	name	kunde	budget
1	DWH	BMW	400000
2	eCommerce	Metro	20000
5	SAP	RAG	50000
...

Integrität

45

- Datentypen
- Schlüssel
- Fremdschlüssel
- Rechte
- Transaktionen

mitarbeiter

p_id	vorname	nachname	alter	proj_id
1	Peter	Müller	32	1
2	Stefanie	Meier	34	1
5	Petra	Weger	28	2
7	Andreas	Zwickel	44	5
...	

projekte

proj_id	name	kunde	budget
1	DWH	BMW	400000
2	eCommerce	Metro	20000
5	SAP	RAG	50000
...

Anfragen

46

- Deklarativ:
 - Nicht „Wie erzeuge ich ein Anfrageergebnis?“
 - Sondern „Was soll im Anfrageergebnis stehen?“
- Sprachlich
 - Nachnamen aller Personen, die in kleinen Projekten arbeiten
- Relationale Algebra
 - $\Pi_{m.nachn.} \sigma_{p.budget < 40000} (\sigma_{m.proj_id = p.proj_id} (\text{mit. } x \text{ proj.}))$
- SQL
 - **SELECT** m.nachname
FROM Mitarbeiter m, Projekte p
WHERE p.budget < 40000
AND m.proj_id = p.proj_id

Anfragebearbeitung – Problem

47

Anfragen

- Anfragen sind deklarativ.
- Anfragen müssen in eine ausführbare (prozedurale) Form transformiert werden.

Ziele

- QEP – prozeduraler Query Execution Plan
- Optimierung (Effizienz)
 - Schnell
 - Wenig Ressourcenverbrauch (CPU, I/O, RAM, Bandbreite)

Anfragebearbeitung

48

Genereller Ablauf

1. Parsen der Anfrage (Syntax)
2. Überprüfen der Elemente (Semantik)
3. Berechnung von möglichen Ausführungsplänen
 - Exponentiell viele
4. Wahl des optimalen Ausführungsplans
 - Regelbasierter Optimierer
 - Kostenbasierter Optimierer
5. Anfrageausführung

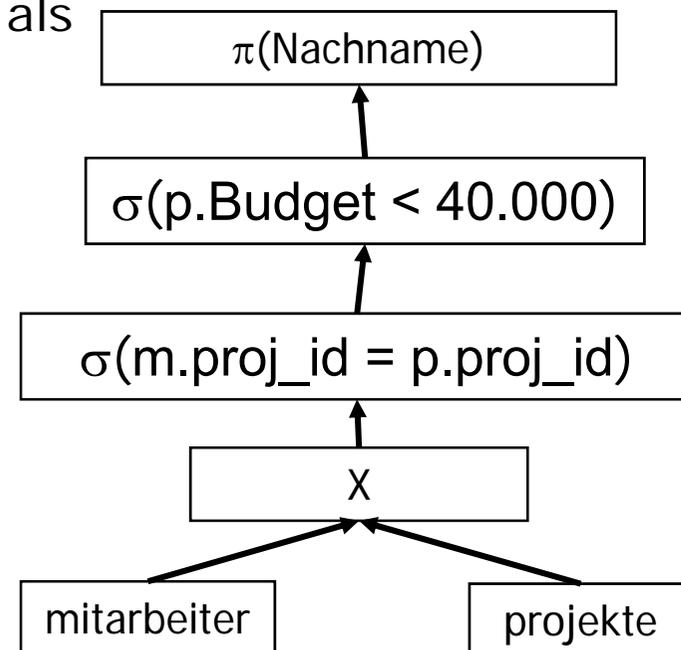
Anfragebearbeitung – Ausführung

49

Interpretation des Ausführungsplans als Baum

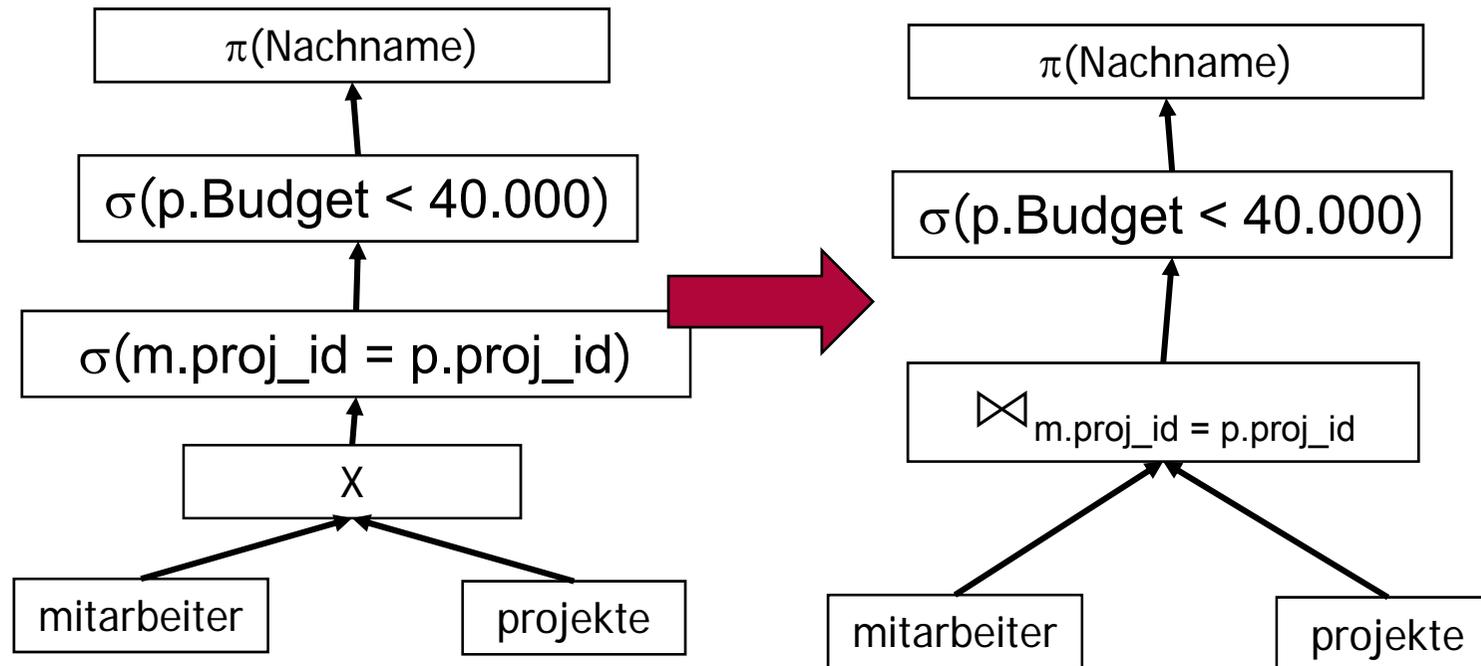
1. Kartesisches Produkt
2. Zwei Selektionen
3. Projektion

Sehr aufwendig!



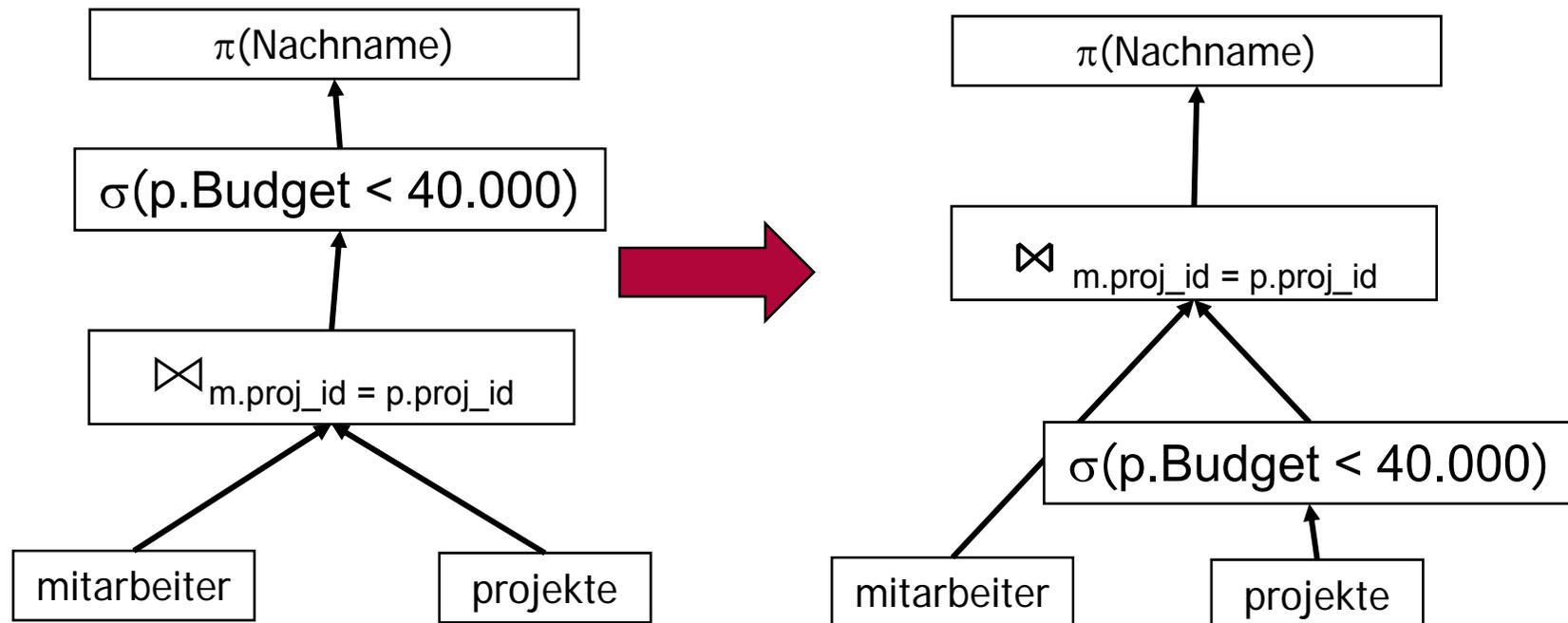
Anfragebearbeitung – Beispiel

50



Anfragebearbeitung – Beispiel

51



Sichten

52

- **CREATE VIEW** MA_MiniProjekte AS (
 SELECT m.nachname
 FROM Mitarbeiter m, Projekte p
 WHERE p.budget < 40000
 AND m.proj_id = p.proj_id)

- **SELECT * FROM** MA_MiniProjekte
- Probleme:
 - Anfrageplanung
 - Materialisierung
 - Updates durch Sichten hindurch

Transaktionsmanagement

53

Problem: Mehrbenutzerbetrieb

- **INSERT INTO** mitarbeiter **VALUES** (Hans, Müller, 36, 5)
- **DELETE FROM** projekte **WHERE** proj_id = 5

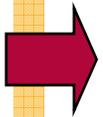
Reihenfolge der Operationen ist nicht unabhängig

- Serialisierbarkeit
- Sperren

Überblick

54

- Vorstellung der Arbeitsgruppe
- Organisatorisches
- Datenbanken und Informationssysteme
- Das Semester an einem Beispiel
- Ausblick auf das Semester



Kurzüberblick

55

- | | | |
|---|--------|---------------------------------------|
| ■ Einführung & Beispiel | | |
| ■ Datenbanksysteme - Historie | | |
| ■ ER-Modellierung | —————→ | ■ Übung 1: ER-Modellierung |
| ■ Relationaler Datenbankentwurf | —————→ | ■ Übung 2: Datenbankentwurf |
| ■ Relationale Algebra | —————→ | ■ Übung 3: Relationale Algebra |
| ■ SQL | —————→ | ■ Übung 4: SQL |
| ■ Constraints & Trigger | | |
| ■ Datenbankprogrammierung | | ■ Übung 5:
Datenbankprogrammierung |
| <hr/> | | |
| ■ Anfragebearbeitung und –
optimierung | —————→ | ■ Übung 6: Anfrageoptimierung |
| ■ Transaktionsmanagement | —————→ | ■ Übung 7: Transaktionen |
| ■ XML und Data Warehouses | | |
| ■ Informationsintegration | | ■ Übung: Klausurvorbereitung |



Die Themen im Einzelnen

56

- Was sind Datenbanken?
 - Motivation, Historie, Datenunabhängigkeit, Einsatzgebiete
- Datenbankentwurf im ER-Modell
 - Entities, Relationships, Kardinalitäten, Diagramme
- Relationaler Datenbankentwurf
 - Relationales Modell, ER -> Relational, Normalformen, Transformationseigenschaften
- Relationale Algebra
 - Kriterien für Anfragesprachen, Operatoren, Transformationen
- SQL
 - SQL DDL, SQL DML, SELECT ... FROM ... WHERE ...

Die Themen

57

- Integrität und Trigger
 - Fremdschlüssel, weitere Bedingungen, Trigger
- Datenbankprogrammierung
 - JDBC, PSM, Embedded SQL, Dynamic SQL
- Sichten und Zugriffskontrolle
 - Sichtenkonzept, Änderungen auf Sichten, Rechtevergabe
- Anfragebearbeitung und -optimierung
 - Anfragebäume, Kostenmodell, Transformationen, Dynamische Programmierung
- Transaktionsverwaltung
 - Mehrbenutzerbetrieb, Serialisierbarkeit, Sperrprotokolle, Fehlerbehandlung, Isolationsebenen in SQL
- Ausblicke
 - Data Warehouses, Data Mining, Informationsintegration, DBS II

Fragen bitte jederzeit!

- In der VL
- Sprechstunde: Dienstags 15 - 16 Uhr
- Email: naumann@hpi.uni-potsdam.de
- Telefon: (0331) 5509 280

Anregungen zur Verbesserung!

The end